working paper
department
of economics

REPUTATION AND EQUILIBRIUM SELECTION
IN GAMES WITH A PATIENT PLAYER

Drew Fudenberg
David Levine

No. 461 July 1987

massachusetts
institute of
technology

50 memorial drive
cambridge, mass. 02139

REPUTATION AND EQUILIBRIUM SELECTION
IN GAMES WITH A PATIENT PLAYER

Drew Fudenberg
David Levine

No. 461 July 1987

REPUTATION AND EQUILIBRIUM SELECTION

IN GAMES WITH A PATIENT PLAYER

Drew Fudenberg, Massachusetts Institute of Technology

David K. Levine, University of California, Los Angeles

July 1987

This paper benefited from conversations with David Kreps and John Riley. National Science Foundation grants SES 85-09697, SES 85-09484, the Sloan Foundation and UCLA Academic Senate provided financial support. This paper was begun when the first author was at the University of California, Berkeley.

1. INTRODUCTION

Consider a game in which a single long-run player faces an infinite sequence of opponents, each of whom play only once. While such a game will often have multiple equilibria, a common intuition is that the "most reasonable" equilibrium is the one which the long-run player most prefers. This paper shows that "reputation effects" provides a foundation for that intuition, and it also identifies an important way in which the intuition must be qualified.

More specifically, imagine that players move simultaneously in each period, and let the "Stackelberg outcome" be the long-run player's most preferred outcome of the stage game under the constraint that each short-run player chooses an action that maximizes his single-period payoff. Now formulate the situation as a game of incomplete information, and imagine that with non-zero probability the long-run player is a "type" who always plays his Stackelberg action. When the discount factor is sufficiently near to one, any Nash equilibrium must give the long-run player almost his Stackelberg payoff. The intuition is the familiar one that the long-run player can choose to play as if he were the type that always plays Stackelberg, and can thus acquire the "reputation" for being a Stackelberg type. This intuition relies on the assumption that the short-run players in fact observe the long-run player's strategy in the stage games, and need not hold in sequential-move games where some actions by the short-run player may prevent the long-run player from acting at all.

Our work builds on that of several previous authors, most directly that of Kreps-Wilson [1982], Milgrom-Roberts [1982], and Fudenberg-Kreps [1987] on reputation effects in the chain-store paradox. These papers considered a long-lived incumbent facing a succession of short-lived entrants, and showed

that if there was a small chance that the incumbent was "tough," it could deter entry by maintaining a reputation for toughness. Our result improves on theirs in several ways, all of which stem from the fact that our results apply to all of the Nash equilibria of the repeated game. First, our results are robust to further small changes in the information structure of the game. The earlier arguments depend on the restriction to sequential equilibria, which as Fudenberg-Kreps-Levine [1987] have argued, is not robust to such changes.

Second, our proof is much simpler, and provides a clearer understanding of the reputation-effects phenomenon. The point is simply that since the short-run players are myopic, they will play as Stackelberg followers in any period they attach a large probability to the long-run player playing like a Stackelberg leader. We use this observation to show that if the long-run player chooses his Stackelberg action in every period, there is an upper bound on the number of times the short-run players can fail to play as Stackelberg followers. This argument is much simpler than the earlier ones, which were obtained by characterizing the sequential equilibria. (The earlier papers did however, obtain characterizations of equilibrium play as well as of the equilibrium payoffs.)

Third, because our proof is simpler, we are able to study a broader class of games. We consider arbitrary specifications of the stage game, as opposed to the special case of the chain store, and we consider a more general form of the incomplete information: Where the earlier papers specified that the long-run player had two or three types, our result covers all games in which the "Stackelberg type" has positive probability. Also, our results extend to non-stationary games in which the long-run player has private information about his payoffs in addition to knowing whether or not

he is a "Stackelberg type."

Our work is also related to that of Kreps-Milgrom-Roberts-Wilson [1982] and Fudenberg-Maskin [1986] on reputation effects in games where all of the players are long-lived, and, more closely, to that of Aumann-Sorin [1987]. Kreps-Milgrom-Roberts-Wilson considered the effects of a specific sort of incomplete information in the finitely-repeated Prisoner's Dilemma, and showed that in all sequential equilibria the players cooperated in all but a few of the periods. Fudenberg-Maskin showed that for any given individually rational payoffs of a repeated game there is a family of "nearby" games of incomplete information each of which has a sequential equilibrium which approximates the given payoffs. Our result does not apply because long-run players need not play short-run best responses. They can tradeoff a loss today for a gain tomorrow. This is not the case with a single long-run player, which is why this paper's results are so different. The work of Aumann-Sorin is closer to ours in developing bounds on payoffs that hold uniformly whenever the incomplete information puts probability on a sufficiently broad class of preferences. To obtain this sort of result with several long-run players, Aumann-Sorin require very strong assumptions: the only stage games considered are those of "pure coordination," and the preferences of the "crazy" types must be represented by "finite automata with finite memory."

2. THE SIMPLE MODEL

We begin with the simplest model of a long run player facing a sequence of opponents. The long-run player, player one, faces an infinite sequence of different short-lived player two's. Each period, player one selects strategy from his strategy set S_1, while that period's player two selects

a strategy from S_2. In this section we assume that players one and two move simultaneously in each period, so that at the end of the period each player knows what strategy his opponent used during that period. Section five considers the complications that arise when the stage game has a nontrivial extensive form. For the time being we will also assume that the S_i are finite sets; Section six considers the technical issues involved when the S_i are allowed to be any compact metric space. Corresponding to the strategy spaces S_i are the spaces Σ_i of mixed strategies.

The unperturbed stage game is a map

$$g: S_1 \times S_2 \to \mathbb{R}^2,$$

which gives each player i's payoff g_i as a function of the realized actions. In an abuse of notation, we let $g(\sigma) = g(\sigma_1, \sigma_2)$ denote the expected payoff corresponding to the mixed strategy σ.

In the unperturbed repeated game $G(\delta)$, the long-run player discounts his expected payoffs using the discount factor δ, $0 < \delta < 1$. Specifically, the sequence of payoffs $g_1^1, \ldots, g_1^t \ldots$ has present value

$$(1-\delta) \sum_{t=1}^{\infty} \delta^t g_1^t$$

Each period's short-run player acts to maximize that period's payoff.

Both long-run and short-run players can observe and condition their play at time t on the entire past history of the game. Let H_t denote the set of possible histories of the game through time t, $H_t = (S_1 \times S_2)^t$. A pure strategy for player one is a sequence of maps $s_1^t\colon H_{t-1} \to S_1$, and a pure strategy for a period-t player two is a function $s_2^t\colon H_{t-1} \to S_2$. Mixed strategies are $\sigma_i^t\colon H_{t-1} \to \Sigma_i$. (Note that if the stage-game corresponded to a non-trivial extensive form, then the realized play in the stage game

would not reveal how player one would have played at all of his information sets, and thus would not reveal player one's choice of a normal-form strategy for that stage.)

This game has been studied by Fudenberg-Kreps-Maskin [1987]. We summarize their results here for the convenience of the reader.

Let $B: \Sigma_1 \to \Sigma_2$ be the correspondence that maps mixed strategies by player one in the stage game g to the best responses of player two. Because the short-run players play only once, in any equilibrium of $G(\delta)$, each period's play must lie in the graph of B.

Fudenberg-Kreps-Maskin prove that a kind of "folk theorem" obtains for games with a single long-run player. Specifically, let V_1 be the set of payoffs for player one attainable in the graph of B when player one is restricted to pure strategies. Let $\underline{v}_1$ be player one's minimax value in the game in which moves by player two that are not best responses to <u>some</u> play by player one are deleted. Then any payoff in V_1 that gives player one at least $\underline{v}_1$ can be attained in a sequential equilibrium if δ is near enough to one.

The point of this paper is to argue that if the game $G(\delta)$ is perturbed to allow for a small amount of incomplete information, then there is a far narrower range of Nash equilibrium payoffs. Roughly speaking, in the perturbed game the long-run player can exploit the possibility of building a reputation to pick out the equilibrium he likes best.

Specifically, define

$$g_1^* = \max_{s_1 \in S_1} \min_{\sigma_2 \in B(s_1)} g_1(s_1, \sigma_2), \tag{1}$$

and let s_1^* satisfy

$$\min_{\sigma_2 \in B(a_1^*)} g_1(s_1^*,\sigma_2) = g_1^*.$$

We call g_1^* player one's Stackelberg payoff and s_1^* a Stackelberg strategy. This differs slightly from the usual formulation, because when player two has several best responses we choose the best response that player one likes least, instead of the one that he prefers. In the usual definition of Stackelberg equilibrium, the follower is assumed to play the best response that the leader most prefers. This is natural in many games with a continuum of strategies, because the leader can make the follower strictly prefer the desired response by making a small change in his strategy. In our finite-action setting, player one cannot break player two's indifference in this way. Thus to have a lower bound on player one's payoff, we need to allow for the case where player two is "spiteful" and chooses the best response that player prefers least. Note also that there may be several Stackelberg strategies. In the next section we provide a condition on the form of the incomplete-information that ensures that player one's worst Nash equilibrium payoff is close to g_1^* when δ is close to one.

Before turning to the incomplete-information game, though, we should clarify the role of mixed strategies for the long-run player. The Fudenberg-Kreps-Maskin "folk theorem" restricts the long-run player to play pure strategies, as does our definition of g_1^*. Consider Figure 1, which displays an example from Fudenberg-Kreps-Maskin. In this game, the long-run player chooses rows and the short-run player chooses columns. If player one mixes with equal weight on both rows, then a best response for player two is to play L, giving player one a payoff of four. On the other hand, restricting player one to pure strategies yields $g_1^* = 1$.

Fudenberg-Kreps-Maskin show that in any equilibrium of the repeated game, whether or not mixed strategies are allowed, the player one payoff is no more than three.

The problem is that if the short-run players believe that player one will mix with equal probabilities each period, player one will always prefer to play up instead of down. If the long-run player could arrange for his choice of mixed strategy to be observed at the end of each period, then the payoff of four could be attained. Having an observable mixed strategy is equivalent to adding a new pure strategy with the same payoffs as the pure strategy. Thus, when mixed strategies are observable, we have

$$g_1^* = \max_{\sigma_1 \in \Sigma_1} \min_{\sigma_2 \in B(\sigma_1)} g_1(\sigma_1, \sigma_2). \tag{2}$$

		L	M	R
Row	U	5,0	1,1	0,-100
	D	3,2	2,1	1,4

Figure 1

When might it be reasonable to suppose that mixed strategies are observable? Imagine that each player can build an infinitely-divisible lottery wheel with a known, independent probability distribution over outcomes. All wheels are spun in each period. Further, each player observes his wheel at the start of each period (before he chooses an action), and his opponents observe that period's spin at the end of the period. (This requires that the player cannot secretly alter the reading of

his wheel.) Then, since any mixed strategy can be implemented by conditioning on the wheel's outcome, the player's choice of mixed strategy is ex-post observable.

A striking fact, which we demonstrate in another paper, Fudenberg and Levine [1987b], is that with the type of perturbation used in this paper, the long run player can do approximately as well as he can with observable mixed strategies, even though mixed strategies cannot be observed. That is, if it is possible to build a reputation for playing a mixed strategy, the long-run player can approximate the corresponding payoff when he is sufficiently patient.

3. THE PERTURBED GAME

This section introduces the perturbed game and gives the first version of our bounds on the long-run player's Nash equilibrium payoff. Section four gives examples to show that this bound cannot in general be improved on, and that there are generally many Nash equilibria.

In the perturbed game, player one knows his own payoff function, but the short-run players do not. We represent their uncertainty about player one's payoffs using Harsanyi's [1967] notion of a game of incomplete information. Player one's payoff is identified with his "type" $\omega \in \Omega$. It is common knowledge that the short-run players have (identical) prior beliefs μ about ω, represented by a probability measure on Ω. In this section we restrict attention to perturbed games with a countable number of types, so that $\Omega = \{\omega_0, \omega_1, \omega_2, \ldots\}$ is a measure space in the obvious way. Section six allows for an uncountable number of types.

The period payoffs in the <u>perturbed game</u> are the same as in the unperturbed game, except that player one's period t payoff $g_1(s_1, s_2, \omega)$ may

now depend on his type. A pure strategy for player one in the perturbed game is a sequence of maps $s_1^t\colon H_{t-1} \times \Omega \to S_1$, specifying his play as a function of the history and his type; a mixed strategy is $\sigma_1^t\colon\ H_{t-1} \times \Omega \to \Sigma_1$. Otherwise the definition of the perturbed game is the same as that of the unperturbed game. The perturbed game is denoted $G(\delta,\mu)$ to emphasize its dependence on the long-run player's discount factor and on the beliefs of the short-run players.

Let type ω_0 have the preferences corresponding to the unperturbed game $G(\delta)$, so that if player one is type ω_0 then

$$g_1(s_1, s_2, \omega_0) = g_1(s_1, s_2).$$

In some cases it will be most natural for $\mu(\omega_0)$ to be near to unity, in others it will not; this is inessential for our argument. We will however require $\mu(\omega_0)$ to be strictly positive. In addition, for any action $s_1 \in S_1$, let the event $\bar{s}_1$ be a type ω such that player one's best strategy in the repeated game is to play s_1 in every period, that is,

$$g_1(s_1, s_2, \bar{s}_1) = g_1(s_1, \bar{s}_2, \bar{s}_1) > g_1(\bar{s}_1, \bar{s}_2, \bar{s}_1)$$
$$\text{for all } \bar{s}_1 \neq s_1 \text{ and } s_2,\ \bar{s}_2.$$

In other words, player one's payoff is independent of player two's action if he plays s_1, and is strictly more than he can get if he plays any other strategy. This clearly implies that playing s_1 in every period is strictly dominant in the repeated game. If s_1 was merely dominant in the stage game, it would not necessarily dominate in the repeated game. In the prisoner's dilemma, for example, defection is dominant at each stage, but certainly is not a good strategy against a tit-for-tat opponent. Clearly if repeated play of s_1 is strictly dominant in the repeated game, Nash equilibrium requires that (if $\bar{s}_1$ and h_t have positive probability)

$s_1^{t+1}(h_t, \bar{s}_1) = s_1$ for all t and almost h_t. The event $\bar{s}_1^*$ means that player one strictly prefers to play the "Stackelberg strategy" s_1^*. We will say that this event corresponds to player one being "the" Stackelberg type.

Since the perturbed game has countably many types and periods, and finitely many actions per type and period, the set of Nash equilibria is a closed non-empty set. This follows from the standard results on the existence of mixed strategy equilibria in finite games, and the limiting results of Fudenberg and Levine [1983, 1986]. Consequently, we may define $\underline{V}_1(\delta, \mu, \omega_0)$ to be the least payoff to a player one of type ω_0 in any Nash equilibrium of the perturbed game $G(\delta, \mu)$. Observe that the minimum is taken over all mixed strategy equilibria, and not merely pure strategy equilibria.

<u>Theorem 1</u>: Assume $\mu(\omega_0) > 0$, and that $\mu(\bar{s}_1^*) \equiv \mu^* > 0$. Then there is a constant $k(\mu^*)$ otherwise independent of (Ω, μ), such that

$$\underline{V}_1(\delta, \mu, \omega_0) \geq \delta^{k(\mu^*)} g_1^* + (1 - \delta^{k(\mu^*)}) \min g_1$$

This says that if the long-run player is patient relative to the prior probability μ^* that he is "tough", then the long-run player can achieve almost his Stackelberg payoff. Moreover, the lower bound on the long-run player's payoff is independent of the preferences of the other types in Ω to which μ assigns positive probability. The condition $\mu^*(\omega_0) > 0$ is necessary for $\underline{V}_1$ to be well defined. We point out in section six that this condition is not in fact essential.

<u>Proof</u>: Fix any (possibly mixed) equilibrium (σ_1^t, σ_2^t) of $G(\delta, \mu)$, and consider the strategy for player one of always playing s_1^*. We will show that player two's equilibrium strategies choose actions outside of $B(s_1^*)$ at most

$k(\mu^*)$ times, where $k(\mu^*)$ is independent of δ and μ.

Fix a history h_t which occurs with positive probability, and such that player one has always played s_1^*. Since $\mu^* > 0$ such histories exist. We shall show that in any such history player two has played outside of $B(s_1^*)$ at most $\bar{k}$ times. To this end, for $1 \le \tau \le t$, let h_τ be the history h_t truncated at time τ. Define

$$\pi_\tau^*(h_{\tau-1}) \equiv \text{Prob}\,[s_1^\tau = s_1^* \,|h_{\tau-1}]$$

to be the probability that (any type of) player one plays s_1^* in period τ conditional on $h_{\tau-1}$. Since $B(s_1^*)$ is the set of best responses to s_1^*, there is a probability $\bar{\pi} < 1$ such that player two will play an action in $B(s_1^*)$ whenever $\pi_\tau^*(h_{\tau-1}) > \bar{\pi}$. Let $\Omega^* = \Omega/\bar{s}_1^*$. We show below that $\pi_\tau^*(h_{\tau-1}) < \bar{\pi}$ is only possible if $\text{Prob}[s_1^\tau = s_1^* \,|h_{\tau-1},\Omega^*]$ is less than $\bar{\pi}$ as well. Thus in any period where a player 2 does not play a best response to s_1^*, he must expect that the "non-Stackelberg" types are unlikely to play s_1^*. Consequently, by playing s_1^* player one can affect a non-negligible increase in his opponent's belief that he is indeed the Stackelberg type.

To make this precise, suppose that in h_t there have been $\hat{k}$ previous periods $\tau \in \hat{T}$ in which $\pi_{\tau+1}^*(h_\tau) \le \bar{\pi}$. Let $h_t^1 = (s_1^*, s_1^*, \ldots, s_1^*)$ be the history of player one's play, and h_t^2 be the history of player two's play; in other words $h_t = (h_t^1, h_t^2)$. Using Bayes Law

$$\text{(3)} \qquad \text{Prob}(\bar{s}_1^* \,|h_t) = \frac{\mu^*\ \text{Prob}(h_t^2 \,|\bar{s}_1^*)}{\mu^*\ \text{Prob}(h_t^2 \,|\bar{s}_1^*) + (1-\mu^*)\ \text{Prob}(h_t \,|\Omega^*)}$$

Our goal is to show that

$$\text{(4)} \qquad \text{Prob}(h_t|\Omega^*) \le \text{Prob}(h_t^2 \,|\bar{s}_1^*)\ \bar{\pi}^{\hat{k}};$$

that is, the prior probability of a "rational" player one playing s_1^* many

times in a row must be small if player two actually failed to play a best respond to s_1^* k many times in a row. The inequalities (3) and (4) clearly imply that

(5) $$\text{Prob}[\bar{s}_1^* \mid h_t] \geq \mu^*/[\mu^* + (1-\mu^*)\,\bar{\pi}^k].$$

Direct computation then shows that if

(6) $$\hat{k} > k = \log[\mu^*(1-\bar{\pi})/(1-\mu^*)]/\log\bar{\pi},$$

then $\text{Prob}[s_1^{t+1} = s_1^* \mid h_t] \geq \text{Prob}[\bar{s}_1^* \mid h_t] > \bar{\pi}$. In other words, if player one has always played s_1^*, and there have been k periods in which this was unexpected in the sense that $\pi^*_{\tau+1}(h_\tau) \leq \bar{\pi}$, then player two's posterior belief that player one will play s_1^*, $\text{Prob}[s_1^{t+1} = s_1^* \mid h_t]$, exceeds $\bar{\pi}$, and player two will optimally respond by playing in $B(s_1^*)$. Consequently, by induction, if player one plays s_1^* forever, he will get less than g_1^* at most k times.

To prove (4), observe that

(7) $$\text{Prob}(h_t \mid \Omega^*) = \text{Prob}(h_t^1, h_t^2 \mid \Omega^*) =$$
$$\text{Prob}(h_t^2 \mid h_t^1, \Omega^*)\ \text{Prob}(h_t^1 \mid \Omega^*).$$

Moreover, given h_t^1, player two's play does not further depend on player one's type, and it follows that $\text{Prob}(h_t^2 \mid h_t^1, \Omega^*) = \text{Prob}(h_t^2 \mid h_t^1, \bar{s}_1^*)$. Further since $h_t^1 = (s_1^*, s_1^*, \ldots, s_1^*)$, $\text{Prob}(h_t^1 \mid \bar{s}_1^*) = 1$, and so $\text{Prob}(h_t^2 \mid h_t^1, \bar{s}_1^*) = \text{Prob}(h_t^2 \mid \bar{s}_1^*)$. Consequently, to demonstrate (4), it suffices to find a bound on $\text{Prob}(h_t^1 \mid \Omega^*)$. We calculate

(8) $$\text{Prob}(h_t^1 \mid \Omega^*) = \text{Prob}(h_t^1, h_{t-1}^1, \ldots, h_1^1 \mid \Omega^*)$$
$$= \prod_{\tau=1}^{t} \text{Prob}(h_\tau^1 \mid h_{\tau-1}^1, \Omega^*)\ .$$

In periods τ outside of $\hat{T}$, when player two plays in $B(s_1^*)$,

$\text{Prob}(h^1_\tau \mid h^1_{\tau-1}, \Omega^*) \le 1$. In periods $\tau \in \hat{T}$, we have from the elementary laws of probability

(9) $$\pi^*_\tau(h_{\tau-1}) = \text{Prob}[s^\tau_1 = s^*_1 \mid h_{\tau-1}]$$
$$= \text{Prob}[\bar{s}^*_1 \mid h_{\tau-1}] + \text{Prob}[s^\tau_1 = s^*_1 \mid h_{\tau-1}, \Omega^*]\ \text{Prob}[\Omega^* \mid h_{\tau-1}].$$

Moreover, $\pi^*_\tau(h_{\tau-1}) \le \bar{\pi}$, since $\tau \in \hat{T}$. From (9) this yields $\text{Prob}(s^\tau_1 = s^*_1 \mid h_{\tau-1}, \Omega^*) \le \bar{\pi}$. Finally, observe that h^1_τ, consisting of s^*_1 played τ times in a row, occurs following $h^1_{\tau-1}$, consisting of s^*_1 played $\tau-1$ times in a row, if and only if $s^\tau_1 = s^*_1$. Consequently

(10) $$\text{Prob}(h^1_\tau \mid h^1_{\tau-1}, \Omega^*) = \text{Prob}(s^\tau_1 = s^*_1 \mid h_{\tau-1}, \Omega^*) \le \bar{\pi} \quad \tau \in \hat{T}.$$

Combining this with (7) and (8) yields (4). ∎

Note that the same proof holds immediately for finitely-repeated games, including the case where $\delta = 1$: If there are enough periods, player one's average payoff cannot be much below the Stackelberg level.

While we have assumed that player two's payoffs are common knowledge, a simple extension allows us to interpret each player two's choice of s_2 as a choice of a strategy mapping his privately-known type into an action. Under this interpretation, player one's Stackelberg strategy s^*_1 is the one that maximizes player one's expected payoff, given that each type of player two chooses a short-run best response. In the chain-store game, for example, if there is a sufficiently high probability that player two is a type which will enter whether or not player one is expected to fight, then player one's Stackelberg action is to acquiesce. (See Milgrom-Roberts [1982] and Fudenberg-Kreps [1987].)

It is also unimportant that player one's payoff be common knowledge in the unperturbed game. Let player one's possible types in the unperturbed

game be $\theta \in \Theta$, with $g_1 = g_1(s_1, s_2, \theta)$ and g_2 independent of θ. Now construct a perturbed game with type space Ω, and assume that $\Theta \subset \Omega$. For each θ, let $s_1^*(\theta)$ be that type's Stackelberg action, that is, its most preferred action in the graph of B, and assume that each of the events $\bar{s}_1^*(\theta)$ has positive prior probability. The proof of theorem 1 shows that each θ can attain his Stackelberg payoff $g_1^*(\theta)$ when δ is near to one. The results given in section six will cover this extension along with several others.

For notational reasons, we assumed that types in the perturbed game have stationary payoffs independent of time and history. As can be seen from the proof of theorem 1, this assumption is irrelevant. Indeed, even the $\bar{s}_1^*$ type(s) can have non-stationary payoffs, provided that playing s_1 is strictly dominant for the whole (infinite-horizon) game.

Rather more strongly, even the unperturbed game need not be stationary. In a non-stationary game we may define the Stackelberg payoff to be the greatest average present value obtainable when the short-run player plays a passive best response in every period. The nature of the best response may, however, depend on either time or history. The argument is similar to that above, except that now the critical probability $\bar{\pi}$ will depend on time and history. Provided only that it is bounded uniformly away from one, that is, $\bar{\pi}_t(h_t) \leq \bar{\pi} < 1$ for all t, and h_t, the proof remains valid. An application along these lines may be found in example 1 below.

4. EXAMPLES

We now present some examples to illustrate the power and the limitations of our result. Example 1 uses some variants of the chain-store game to illustrate the advantage of our technique of proof: we obtain

asymptotic results without the need to explicitly solve for the sequential equilibria. Example 2 shows that the equilibria need not be unique. Example 3 raises a deeper concern: our result in general becomes much weaker if the stage game is not a simultaneous move. This may seem surprising, because the chain-store game considered by Kreps-Wilson [1982] and Milstrom-Roberts [1982] has sequential moves and not simultaneous ones. As we show in section five, their positive results are due to the special nature of the payoffs that they considered.

Example 1: Consider the following version of Selten's [1977] chain-store game. Each period, a short-run entrant decides whether or not to enter and the long-run incumbent chooses whether to fight or to acquiesce. For conformity with the assumptions of theorem 1, we assume that these choices are made simultaneously, and that at the end of the period the incumbent's choice is revealed, whether or not entry in fact occurred. The entrants differ in two ways. First, each period's entrant is either "strong" or "weak". Strong entrants always enter, weak ones have payoffs described below. Each period's entrant is weak with probability p, independent of the others, and being strong or weak is private information. Second, each entrant is one of three "sizes," big, medium, or small. Each entrant has probability z_b of being big, z_m of being medium-sized, and z_s of being small, and the entrants' sizes are public information. To preserve a stationary structure, we imagine that the incumbent learns the period t entrant's size at the start of period t. Each weak entrant receives one if it stays out, zero if there is a fight, and two if it enters and the incumbent acquiesces. Thus weak entrants enter if the probability that they will be fought is less than one half. The incumbent receives two if it

acquiesces and four if no entry occurs. The incumbent receives $3c/2$ if it has to fight a small entrant, c if it has to fight a medium one, and zero if it has to fight a large one. Thus in the unperturbed game, there is an equilibrium in which all entrants enter and the incumbent acquiesces to all entry. The previous papers on the chain-store game had only one size of entrant, and so the possible "reputations" the incumbent would want (that is, the possible Stackelberg actions) to establish is for always fighting or for always acquiescing.

To find the Stackelberg strategy here, we compute the difference in payoffs between fighting and acquiescing, given that the entrants play their best response:

<u>entrant size</u>	<u>gain to fighting</u>
big	$4p-2$
medium	$(4-c)p - (2-c)$
small	$[(8-3c)p - (4-3c)]/2$

If, for example, $c=1$ and $1/3 < p < 1/2$, the Stackelberg strategy is to fight the small and medium-sized entrants, and acquiesce to the large ones. If the only types of the incumbent are the original one ω_0 and a type ω_1 that will always fight, then this reputation need not be credible: the first time the incumbent concedes to a large entrant he reveals that he is not type ω_1. But if there is a non-zero prior probability that the incumbent is a type that acquiesces only to large entrants, then the incumbent can develop a reputation for playing in this way.

We can modify example 1 in several ways. First, let us consider how our result extends to nonstationary environments. Imagine that there are only medium size firms and that $p = 0.01$ in odd-numbered periods, and $p = 0.60$ in even-numbered periods. Player one's <u>constant</u> strategy is then

to always acquiesce, because the cost (-0.97) of fighting in the odd periods outweighs the potential gain (0.80) from entry deterrence in the even ones. However, if we allow player one the opportunity to maintain a reputation for nonstationary play, he can do much better. If the prior distribution on Ω assigns positive probability to player one fighting in even periods and acquiescing in the odd ones, the proof of theorem 1 extends in the obvious way.

Next, imagine that the incumbent plays the chain-store game against two simultaneous sequences of entrants: each period the incumbent faces one entrant in market A and another in market B. Each entrant's payoff depends only on play in its own market, but all entrants observe previous play in both markets, so that the markets are "informationally linked" in the sense of Fudenberg-Kreps [1987]. Entrants in market A are "weak" with probability P_A, each independent of the others, while entrants in market B are strong with probability P_B. Fudenberg-Kreps, in a similar but more complex setting, assume that the entrants believe that the incumbent either has the payoffs of the unperturbed version of example 1 or will fight all comers, so that once the incumbent quits in one market he is revealed as weak in both of them. However, if the perturbed game puts positive weight on the incumbent only fighting in one of the market, he is free to develop that reputation.

<u>Example 2</u>: Consider the 2×2 game.

	L	R
U	1,1	0,0
D	0,0	10,10

Figure 2

If $\mu(\omega=\omega_0)$ is near to one, then, regardless of δ, the game $G(\delta,\mu)$ has several sequential equilibria, all of which satisfy our bound. For concreteness, suppose that $0 < \epsilon \le 1/11$, and that μ puts weight $1-\epsilon$ on ω_0, weight ϵ on $\bar{s}_1^*$ (the player who always plays D), and no weight on any other types. One equilibrium of this game has the short-run players playing R regardless of history, and both types of player one playing D regardless of history. Another equilibrium has the first period's player two playing L, and type ω_0 player one playing U in the first period. Play from the second period on matches that in the first equilibrium. Neither player one nor the first period's player two has an incentive to deviate, as the first-period actions are a static Nash equilibrium, and not deviating gives player one his highest possible payoff from period two on. There is no point in player one initially imitating the $\bar{s}_1^*$ type, since he will do just as well beginning next period anyway.

Yet a third equilibrium has type ω_0 reverting to L and player two to D after period T. This is an equilibrium provided T is large enough and δ small enough that $\delta^{T+1}/(1-\delta) \le 1/9$. In this case, the loss to player one by playing U in period one (of one) exceeds the potential gain from convincing two that $\omega = \alpha_1^*$ (of $9\delta^{T+1}/(1-\delta)$. Moreover, once one has revealed in period one (by playing U) that he is not type $\bar{s}_1^*$, he cannot later build a reputation for being this type.

<u>Example 3</u>: This example has the same extensive form as the sequential-move version of the chain store game, but different payoffs. Player two begins by choosing whether or not to purchase a good from player one. If he chooses, both players receive zero. If he buys player one can produce high quality or low. High quality gives each player one, while low quality gives

player one a payoff of three and player two a payoff of zero (see Figure 3). The Stackelberg outcome here is for player one to "promise" to choose high quality, so that all the customers will come in. Thus if theorem 1 extended to this game it would say that if there is a positive prior probability μ^* that player one is a type that always produces high quality, then if δ is near enough to unity, player one's payoff must be very close to one in any Nash equilibrium. However, this extension is clearly false. Take $\mu^* = .01$ and $\mu(\omega_0) = .99$, and specify that the "sane" type ω^0 always produces low quality, so the player two's never buy. Given their pessimistic beliefs, the player twos are correct to not buy and so player one does not have the opportunity to demonstrate that he will produce high quality. If the player two's were a single long-run player, they might be tempted to purchase once or twice to gather information, but myopic player twos will not make this investment. Thus we see that for general stage games it is not true that player one can ensure almost his Stackelberg payoff.

We have two responses to the problem posed by example 3. The first is to follow Fudenberg-Maskin and Fudenberg-Kreps-Levine and examine perturbations with the property that every information set of the stage game is reached with positive probability. (See Fudenberg-Kreps-Levine for an explanation of the perturbations involved.) In this case once again we know that by playing his Stackelberg action s_1^* in every period, player one can eventually force play to the Stackelberg outcome.

Our second response is developed in the next section, which gives a lower bound on player one's payoff that holds for general games.

Before developing that argument, let us explain why the problem raised in example 3 does not arise in the chain store paradox (Figure 4). There, the one action the entrant could take that "hid" the incumbent's strategy

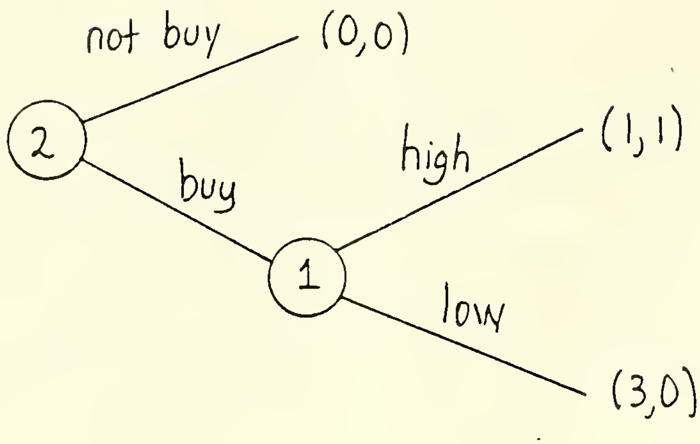

Figure 3: A Quality Game

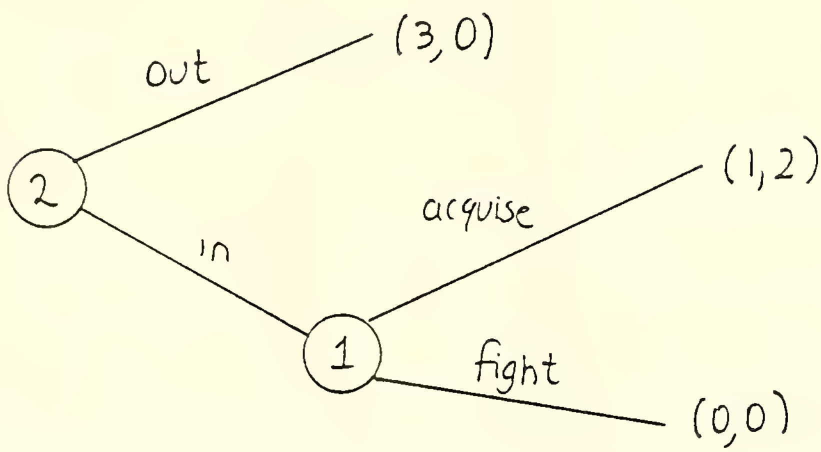

Figure 4: The Classical Chain Store Paradox

was precisely the action corresponding to the Stackelberg outcome: whenever the agent did not play like a Stackelberg follower, the incumbent had a method of demonstrating that the entrant's play was "mistaken". The proof of theorem 1 invoked Bayesian inference only in those periods where s_2^t does not belong to $B(s_1^*)$; whether the incumbent's strategy is revealed in periods where $s_2^t \in B(s_1^*)$ is irrelevant.

5. GENERAL FINITE STATE GAMES

This section develops two extensions of the basic argument of theorem 1. Section 5a treats the case with several interacting short-run players in each period, and section 5b handles general but finite two-player stage games. We defer the technical complications posed by uncountably many actions and types to section six.

5a. Several Short-Run Players

Imagine now that the stage game is a finite n-player simultaneous move game

$$g: S_1 \times S_2 \times \ldots \times S_n \to \mathbb{R}^n,$$

with player one the only long-run player. Since each of the short-run players must play a short-run best response, in equilibrium each period's outcome must lie in the best response sets of all of the short-run players, that is, it must be a Nash equilibrium in the $(n-1)$ player game induced by fixing a (possibly mixed) action by player one. Let $B: \Sigma_1 \rightrightarrows \Sigma_2 \times \ldots \times \Sigma_n$ be this Nash correspondence. Note that this definition of $B(s_1)$ agrees with our previous one for the case of a single short-run player. The corresponding definition of g_1^* is thus

$$g_1^* \equiv \max_{s_1} \min_{\sigma_{-1} \in B(s_1)} g_1(s_1, \sigma_{-1}).$$

As before, let s_1^* be a Stackelberg action for player one, that is, an action that attains g_1^*.

This situation is much the same as with a single short-run player, and as one would expect, the approach of theorem 1 can be readily extended. There is only one minor complication: in the proof of theorem 1 we argued that since the set $B(s_1^*)$ contained all the best responses to s_1^*, then there was a probability $\bar{\pi} < 1$ such that if player one was expected to play s_1^* with probability exceeding $\bar{\pi}$, player two would choose an action in $B(s_1^*)$. With several short-run players, the Nash correspondence $B(\cdot)$ need not be constant in the neighborhood of s_1^*, but it is still upper hemi-continuous. Thus, when player one plays s_1^* and his opponent plays a Nash equilibrium for some σ_1 near to s_1^*, the <u>lowest</u> player one's payoff can be approximately g_1^*.

As before, let ω_0 be the type whose payoffs are as in the unperturbed game, and $\bar{s}_1^*$ be the event that player one has "s_1^* forever" as a best strategy.

<u>Theorem 2</u>: Fix a game $G(\delta)$, with several short-run players and consider a perturbed version $G(\delta,\mu)$. Assume $\mu(\omega_0) > 0$, and that $\mu(\bar{s}_1^*) \equiv \mu^* > 0$. Then for all $\epsilon > 0$, there is a $\underline{\delta} < 1$ such that for all $\delta \in (\underline{\delta},1)$

$$\underline{V}_1(\delta,\mu,\omega_0) \geq (1-\epsilon)g_1^* + \epsilon \min g_1.$$

Moreover, $\underline{\delta}$ depends on μ only through μ^*.

<u>Proof</u>: For any $\pi \in (0,1]$ let $B(\pi,s_1^*)$ be the set of all mutual best responses by the short-run players to each other and to any strategy for player one that puts probability at least π on s_1^*:

$$B(\pi, s_1^*) = \bigcup_{\sigma_1 | \sigma_1(s_1^*) > \pi} B(\sigma_1),$$

and let

$$d(\pi) = \inf_{\sigma_{-1} \in B(\pi, s_1^*)} g_1(s_1^*, \sigma_{-1})$$

be the function that bounds how low player one's payoff can be when he plays s_1^* and the short-run players choose actions in $B(\pi, s_1^*)$. Clearly $d(1) = g_1^*$; and $d(\pi)$ is non-decreasing. We further claim that $d(\pi)$ is continuous in the neighborhood of $\pi=1$. To see this, suppose to the contrary that there is an $\epsilon > 0$ and a sequence $\pi^n \to 1$ such that for all n, $d(\pi^n) < g_1^* - \epsilon$. Then there is a sequence $\sigma_{-1}^n \in B(\pi^n, s_1^*)$ such that $g_1(s_1^*, \sigma_{-1}^n) < g_1^* - \epsilon$. Extracting a convergent subsequence from the σ_{-1}^n, and using the upper hemicontinuity of $B(\cdot)$, we conclude that there is a $\bar{\sigma}_{-1} \in B(s_1^*)$ with $g_1(s_1^*, \bar{\sigma}_{-1}) < g_1^* - \epsilon$, which contradicts the definition of g_1^*.

To prove the theorem, fix an $\epsilon > 0$, and choose π such that $d(\pi) > (1-\epsilon)g_1^*$. As in the proof of theorem 1, if player one chooses the strategy of always playing s_1^*, there is a bound $k(\mu^*, \pi)$, independent of δ, on how many times the short-run players can choose actions that are not in $B(\pi, 1)$. Finally, take δ large enough that δ^k exceeds $(1-\epsilon)$. ∎

Our results use the hypothesis that player one's opponents are short-lived only to ensure that they always play myopically. Thus, theorem 2 extends to games where a "large" player one faces a continuum of "small" opponents, with the (non-innocuous) assumption that no player can observe the actions of a set of opponents of measure zero. This makes precise a sense in which being infinitely larger than one's opponents is the same as

being infinitely more patient than they are.

The large- and small-players case differs from the long- and short-lived one in that our results for the latter hold for any discount factor, while for the former they hold only in the continuum of players limit. An exact analogy between the cases would require that there be a bound on player one's payoffs when he faces small but not infinitesimal opponents, but this is not possible without further assumptions. The difficulty is that when players are small but not infinitesimal, they can have a large influence on equilibrium play. This is why the assumption that measure-zero deviates are ignored is not innocuous. This is discussed in Fudenberg-Levine [1987a].

5b. General Deterministic Stage Games

Here we take up the point raised by example 3. If the stage-game is not simultaneous-move, the long-run player may not have the opportunity to develop the reputation he would desire. For simplicity we return to the case of a single short-run player, player two. Let the stage-game be a finite extensive form of perfect recall without moves by Nature. As in example 3, the play of the stage game need not reveal player one's choice of normal-form strategy s_1. However, when both players use pure strategies the information revealed about player one's play is deterministic. Let $O(s_1,s_2) \subset A_1$ denote the strategies s_1' of player one such that (s_1',s_2) leads to the same terminal node as (s_1,s_2). We will say that these strategies are observationally equivalent meaning that the player two's do not observe player one's past normal-form strategies but only the realized outcomes. In example 3, player one's Stackelberg action s_1^* was to produce high quality but, given that player two chooses not to risk dealing with him, player one had no way of establishing a reputation for Stackelberg

play. The problem was that while "do not buy" was not a best response to s_1^*, do not buy is a best response to "low quality" and high and low quality are observationally equivalent when player two does not purchase.

This suggests the following generalization of theorem 1: For any s_1, always playing s_1 should eventually force player two to play a strategy s_2 which is a best response to an s_1' in $O(s_1,s_2)$. That is, for each s_1 let $W(s_1)$ satisfy

$$W(s_1) = \{s_2 | \exists \sigma_1' \text{ with } \operatorname{supp} \sigma_1' \in O(s_1,s_2) \text{ such that } s_2 \in B(\sigma_1')\}.$$

In other words, $W(s_1)$ is the set of best responses for player two, to beliefs about player one's strategy, that are consistent with the information revealed when that response is played. Then if δ is near to one, player one should be able to ensure approximately

$$g_1^* = \max_{s_1} \min_{s_2 \in W(s_1)} g(s_1,s_2).$$

As before, let s_1^* be a strategy that attains g_1^*, and let $\bar{s}_1^*$ be the event that player one's best strategy in $G(\delta,\mu)$ is to always play s_1^*. Note that if the stage-game is simultaneous move, $W(s_1) = B(s_1)$, and the definitions of g_1^* and s_1^* reduce to those we gave earlier.

<u>Theorem 3</u>: Let g be as described above. Assume $\mu(\omega_0) > 0$ and that $\mu(\bar{s}_1^*) = \mu^* > 0$. Then there is a constant $k(\mu^*)$, independent of μ, such that

$$\underline{V}_1(\delta,\mu,\omega_0) \geq \delta^{k(\mu^*)} g_1^* + (1-\delta^{k(\mu^*)}) \min g_1.$$

Before giving the proof, let us observe that this result, while not as strong as the assertion in theorem 1 that player one can pick out his preferred payoff in the graph of B, does suffice to prove that player one

can develop a reputation for "toughness" in the sequential-move version of the chain store game. Consider the extensive form in Figure 4 above. In this game $B(\text{fight}) = \{\text{out}\}$ and $B(\text{acquiesce}) = \{\text{in}\}$. Also, $O(\text{fight},\text{out}) = O(\text{acquiesce},\text{out}) = \{\text{acquiesce, fight}\}$, while $O(\text{fight},\text{in}) = \{\text{fight}\}$ and $O(\text{acquiesce},\text{in}) = \{\text{acquiesce}\}$.

First, we argue that $W(\text{fight}) = B(\text{fight})$. To see this observe that $W(\text{fight}) \supseteq B(\text{fight}) = \{\text{out}\}$. Moreover, "in" is not a best response to "fight", and "acquiesce" is not observationally equivalent to "fight" when player two plays "in". Consequently, no strategy placing positive weight on "in" is in $W(\text{fight})$.

Finally, since player one's Stackelberg action with observable strategies is fight, and $W(\text{fight}) = B(\text{fight})$, the fact that only player one's realized actions, and not his strategy, is observable does not lower our bound on player one's payoff.

<u>Proof of Theorem 3</u>: Once again we fix an equilibrium (σ_1^t,σ_2^t) and consider the strategy for player one of always playing s_1^*. Let $\pi(h_t)$ be the probability distribution over S_1 that player two expects player one to use in period t. (This is computed from player two's initial beliefs μ, the observed history h_t, and player one's equilibrium strategy in the usual way.) If $s_2 \notin W(s_1^*)$, then there is a $\bar{\pi}(s_2)$ such that s_2 is not a best response to any σ, with $\sigma_1^*(s_1^*,s_2) \geq \bar{\pi}$. Let $\bar{\pi} \equiv \max \bar{\pi}(s_2)$. Each time player two plays an s_2^t outside of $W(s_1^*)$, the observed outcome will be one that had prior probability less than $\bar{\pi}$. Thus, as in the proof of theorem 1, the probability that player one is the type that plays s_1^* increases a non-negligible amount. The rest of the proof is the same as before. ∎

6. GAMES WITH A CONTINUUM OF STRATEGIES

We turn attention now to the case where players have a continuum of strategies in each period. Two complications arise in the analysis. First, it is no longer true that merely because the short run player puts a large probability weight on the Stackelberg strategy, he must play a best response to it. However, he must play an ϵ-best response, and this is sufficient for our purposes. Second, it is not sensible to suppose that *a priori* the short run player places positive weight on the Stackelberg strategy: instead we assume that all neighborhoods of the Stackelberg strategy have positive probability. Nash equilibrium, then, does not require the short run player to optimize against the Stackelberg type, since that type occurs with probability zero. Instead, we must work with a sequence of types that converge to the Stackelberg type.

We return to the basic simultaneous move model. Our description of the perturbed and unperturbed game is unchanged with two exceptions. First, S_1 and S_2 are now assumed to be compact metric spaces, rather than finite sets, and Ω may be an arbitrary measure space. Second, the payoff maps $g_1: S_1 \times S_2 \times \Omega \to \mathbb{R}$ and $g_2: S_1 \times S_2 \to \mathbb{R}$ are assumed to be continuous on $S_1 \times S_2$. The definition of Stackelberg strategies, payoffs and types remains unchanged.

In order to deal with the continuum case, we need to consider ϵ-best responses by player two. Define $B_\epsilon: \Sigma_1 \to \Sigma_2$ to be the correspondence that maps mixed strategies by player one in the stage game g to ϵ-best responses of player two. That is, if $\sigma_2 \in B_\epsilon(\sigma_1)$, then $g_2(\sigma_1,\sigma_2) + \epsilon \geq g_2(\sigma_1,\sigma_2')$ for all $\sigma_2' \in \Sigma_2$. Let d denote the distance metric. Recall the Stackelberg strategy $s_1^*(\omega)$ depends on the type of player 1, $\omega \in \Omega$. We define a corresponding version of the Stackelberg

payoff

$$g_1^*(\epsilon,\omega) = \min_{\substack{\sigma_2 \in B_\epsilon(s_1^*(\omega)) \\ d(s_1,s_1^*(\omega)) \le \epsilon}} g_1(s_1,\sigma_2,\omega).$$

In other words, we allow ϵ-best responses to strategies that differ from $s_1^*(\omega)$ by up to ϵ. We also emphasize the dependence of g_1^* on ω. If for some ω there are several Stackelberg strategies $s_1^*(\omega)$, each will generally yield a different function $g_1^*(\epsilon,\omega)$. However, all version have one key feature in common.

<u>Lemma 4</u>: $\lim_{\epsilon \to 0} g_1^*(\epsilon,\omega) = g_1^*(\omega)$.

<u>Proof</u>: Fix ω. If the lemma fails, there exists a sequence $\epsilon^n \to 0$, $s_1^n \to s_1^*$ and ϵ^n-best responses σ_2^n to s_1^n such that $\lim g_1(s_1^n,\sigma_2^n) < g_1^*$. Since s_2 is compact, Σ_2 is weakly compact, and we may assume $\sigma_2^n \to \sigma_2$. Since g_1 is weakly continuous $\lim g_1(s_1^n,\sigma_2^n) = g_1(s_1^*,\sigma_2) < g_1^*$. From the definition of g_1^* it is clear that σ_2 cannot be a best response to s_2^*; that is, there exists σ_2^* with $g_2(s_1^*,\sigma_2^*) \ge g_2(s_1^*,\sigma_2) + \epsilon$ for some $\epsilon > 0$. However, since g_2 is weakly continuous, we have $g_2(s_1^n,\sigma_2^*) \to g_2(s_1^*,\sigma_2^*)$ and $g_2(s_1^n,\sigma_2^n) \to g_2(s_1^*,\sigma_2)$. This implies

$$g_2(s_1^n,\sigma_2^*) > g_2(s_1^n,\sigma_2^n) + \epsilon/2 \ge g_2(s_1^n,\sigma_2^n) + \epsilon^n,$$

contradicting the fact that σ_2^n is an ϵ^n-best response to s_1^n. ∎

To prove that as $\delta \to 1$, the long run player can get nearly the Stackelberg payoff, must must describe our assumption on μ, the distribution over types. Consider, for each ω, the set of strategies s_1 with $d(s_1,s_1^*(\omega)) \le \bar{d}$. Corresponding to these are types $\bar{s}_1$ who surely play the corresponding strategy. Define $\mu^*(\bar{d},\omega)$ to be the probability assigned to

these types by μ. We can prove

<u>Theorem 5</u>: For a set Ω_0 occurring with positive probability assume that $\inf_{\omega_0 \in \Omega_0} \mu^*(\bar{d},\omega_0) > 0$ for $\bar{d} > 0$. Then for almost all $\omega_0 \in \Omega_0$ and all $\epsilon > 0$ there exists a $\underline{\delta} < 1$ such that for all $\delta \in (\underline{\delta},1)$

$$\underline{V}_1(\delta,\mu,\omega_0) \geq (1-\epsilon)g_1^*(\omega_0) + \epsilon \min g_1.$$

Moreover, $\underline{\delta}$ depends on μ only through $\mu^*(\cdot)$.

<u>Proof</u>: The reason that the theorem holds only for almost all $\omega_0 \in \Omega_0$ is that $\underline{V}_1(\delta,\mu,\omega_0)$, a conditional expectation, is only defined and need only be chosen optimally by player one almost everywhere. Fixing such an ω_0, the proof is essentially the same as that of theorem 1. In the proof of that theorem, we made use of the fact that there is a history h_t that occurs with positive probability and such that player one has always played s_1^*, to conclude that player two must respond optimally given the posterior probabilities based on h_t. Choosing a small number ϵ' notice that the probability of $\bar{s}_1$ with $d(s_1,s_1^*) < \epsilon'$ is positive. It follows that for some s, with $d(s_1,s_1^*) < \epsilon'$, player two must respond optimally following the history h_t that results when s_1 is played repeatedly. In particular, since of the types $\bar{s}_1'$ with $d(s_1',s_1) < \epsilon'$, it is almost surely true in equilibrium that only type $\bar{s}_1$ will actually play s_1, we can choose s_1 so that after observing s_1 played, the type $\bar{s}_1$ will have positive conditional probability. In the proof of theorem 1, we also made use of the fact that for some $\bar{\pi} < 1$ player two will play an action in $B(s_1^*)$ whenever $\pi_\tau^*(h_{\tau-1}) > \bar{\pi}$. The corresponding fact here is that for each ϵ', there is some $\bar{\pi} < 1$ such that player two will play an action in $B_\epsilon(s_1)$ whenever $\pi_\tau^*(h_{\tau-1}) > \bar{\pi}$. With these adaptions, the proof of theorem

1 shows that

$$\underline{V}_1(\delta,\mu,\omega_0) \geq (1-\delta^k)g_1^*(\epsilon',\omega_0) + \delta^k \min g_1.$$

where k depends only on ϵ'. The theorem now follows from lemma 4. ∎

Let us conclude by observing that theorem 5 can be extended along the lines of Section 5b to cover general stage games. This simply involves introducing the set $W_\epsilon(s_1)$ of strategies of player two that are a ϵ-best response to beliefs that are consistent with the information revealed when that response is played. Then $s_1^*(\omega)$ and $g_1^*(\omega)$ are defined in the obvious way, and we proceed as above.

This extension allows us to treat the sequential play of the Kreps-Wilson [1982] two-sided predation game, which was analyzed in Fudenberg-Kreps [1987]. The stage game is played on the interval [0,1], with each player choosing a time to concede if the other player was still fighting. If the player two's were unlikely to be "tough", the Stackelberg action for player one is to commit to fight until the finish $(t=1)$, which induces the "weak" player two's to concede immediately. This game is not a simultaneous move: if the first player two concedes immediately, the others will not learn how long player one would have been willing to fight him. However, as in the simple predation game of example 2, this does not pose additional complications, because whenever a player two does not play as a Stackelberg follower player one will have a chance to demonstrate that he is "tough". That is, $B_\epsilon(t=1) = W_\epsilon(t=1)$. We conclude that if player one is patient he can do almost as well as if he could commit himself to never conceding.

REFERENCES

Aumann, R. and S. Sorin [1987], "Cooperation and Bounded Rationality."

Fudenberg, D. and D. Kreps [1987], "Reputation and Simultaneous Opponents," forthcoming in Review of Economic Studies.

Fudenberg, D., D. Kreps and D. Levine [1986], "On the Robustness of Equilibrium Refinements," forthcoming, Journal of Economic Theory.

Fudenberg, D. and D. Levine [1983], "Subgame-Perfect Equilibria of Finite and Infinite Horizon Games," Journal of Economic Theory, 31, 251-268.

——— [1986], "Limit Games and Limit Equilibria," Journal of Economic Theory, 38, 261-279.

——— [1987a], "Open-Loop and Closed-Loop Equilibria in Dynamic Games with Many Players," forthcoming in Journal of Economic Theory.

——— [1987b], "Reputation, Mixed Strategies, and Martingales," UCLA.

Fudenberg, D. and E. Maskin [1986], "The Folk Theorem in Repeated Games with Discounting or with Incomplete Information," Econometrica, 54, 533-554.

Fudenberg, D., D. Kreps and E. Maskin [1987], "Repeated Games With Long- and Short-Lived Players," UC Berkeley.

Harsanyi, J. [1967], "Games with Incomplete Information Played by Bayesian Players," Management Science, 14, 159-182, 320-334.

Kreps, D., P. Milgrom, J. Roberts, and C. Wilson [1982], "Rational Cooperation in the Finitely Repeated Prisoners' Dilemma," Journal of Economic Theory, 27, 245-252, 486-502.

Kreps, D. and C. Wilson [1982], "Reputation and Imperfect Information," Journal of Economic Theory, 27, 253-279.

Milgrom, P. and J. Roberts [1982], "Predation, Reputation and Entry Deterence," Econometrica, 50, 443-460.

Selten, R. [1977], "The Chain-Store Paradox," Theory and Decision, 9, 127-159.